REPORTS

Definiteness and Semidefiniteness of Quadratic Forms Revisited

Yves Chabrillac and J.-P. Crouzeix
Département de Mathématiques Appliquées
Université de Clermont II
63170 Aubière, France

ABSTRACT

It is shown how the Schur complement theory can be used for the derivation of criteria for the definiteness (or semidefiniteness) of the restriction of a quadratic form to the null space of a matrix.

1. INTRODUCTION

In this paper, we present a survey of different criteria for the positive definiteness and positive semidefiniteness of the restriction of a quadratic form on R^n to the null space of a $p \times n$ matrix. This question is of interest in second-order local minimality conditions for equality constrained minimization problems and has been the object of considerable attention in the last few decades. Among major contributions we mention those of Hancock [7], Mann [11], Samuelson [12], Debreu [5], Bellman [1], and Hestenes [10]. The particular case $p = 1$ has been specially investigated by Crouzeix and Ferland [4] in connection with second-order characterizations of quasiconvexity and pseudoconvexity (see also Schaible [14]).

We will show that the Schur complement theory can be used to unify the derivation of criteria, some of them being new.

2. NOTATION AND MATHEMATICAL BACKGROUND

Throughout the paper, A denotes an $n \times n$ real symmetric matrix, and B and $n \times p$ real matrix of rank q such that $0 \leqslant q \leqslant \max(n, p)$. Our purpose is to unify the treatment of conditions which are equivalent to

(E_1) $x^TAx > 0$ for all $x \in R^n$ such that $B^Tx = 0$ and $x \neq 0$, and
(C_1) $x^TAx \geqslant 0$ for all $x \in R^n$ such that $B^Tx = 0$.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

Of course, condition (E_1) implies condition (C_1). The $(n+p)\times(n+p)$ matrix

$$\mathscr{A} = \begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}$$

is referred as the *bordered matrix* associated with A and B.

Given any real symmetric $r \times r$ matrix M, the *inertia* of M, denoted by $\operatorname{In}(M)$, is the triple $(\pi(M), \nu(M), \delta(M))$ consisting of the numbers of positive, negative, and zero eigenvalues of M, so that $\pi(M)+\nu(M)+\delta(M)=r$. If P is any nonsingular real $r \times r$ matrix, then $\operatorname{In}(M)=\operatorname{In}(P^TMP)$, and if N is any real symmetric positive semidefinite matrix, then the functions of one real variable which associate with $t \in R$ the numbers $\pi(M+tN)$, $\nu(M+tN)$, $\delta(M+tN)$ are respectively nondecreasing and lower semicontinuous, nonincreasing and lower semicontinuous, and upper semicontinuous.

Given the partitioned matrix

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{pmatrix}$$

where M_{11} and M_{22} are square symmetric matrices of orders m_1 and m_2 respectively, M_{11} is nonsingular, and M_{12} is a $m_1 \times m_2$ matrix, then the Schur complement of M_{11} in M is $M_{22} - M_{12}^T M_{11}^{-1} M_{12}$ and denoted by M/M_{11}. Then

$$\det(M) = \det(M_{11})\det(M/M_{11}).$$

Another very nice property of the Schur complement expresses the inertia of M in terms of the inertias of matrices M_{11} and M/M_{11}. It says that

$$\operatorname{In}(M) = \operatorname{In}(M_{11}) + \operatorname{In}(M/M_{11}),$$

the addition on the triples being the usual addition of vectors.

Results on the Schur complement can be mainly found in Haynsworth [8], Haynsworth and Ostrowski [9], and Cottle [3]. For a comprehensive survey see Ouellette [11].

3. CONDITIONS IN TERMS OF THE INERTIA OF THE BORDERED MATRIX

First, notice that the inertia of the bordered matrix $\mathscr{A}$ is not changed when a permutation is performed on the columns of the matrix B. Hence, since B is of rank q, we can assume that the first q columns of B are linearly independent and that the last $p-q$ columns are linear combinations of the first q columns. Thus B can be partitioned as $B=(B_1, B_1C)$ where B_1 is an $n \times q$ matrix of rank q and C a $q \times (p-q)$ matrix. Now,

$$\mathscr{A} = \begin{pmatrix} A & B_1 & B_1C \\ B_1^T & 0 & 0 \\ C^TB_1^T & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} I_n & 0 & 0 \\ 0 & I_q & 0 \\ 0 & C^T & I_{p-q} \end{pmatrix} \begin{pmatrix} A & B_1 & 0 \\ B_1^T & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I_n & 0 & 0 \\ 0 & I_q & C \\ 0 & 0 & I_{p-q} \end{pmatrix}.$$

It follows that $\mathrm{In}(\mathscr{A}) = \mathrm{In}(\mathscr{A}_1) + (0,0,p-q)$, where

$$\mathscr{A}_1 = \begin{pmatrix} A & B_1 \\ B_1^T & 0 \end{pmatrix}.$$

Now, since B_1 is of rank q, there exists a real nonsingular $n \times n$ matrix P such that the first $n-q$ rows of the matrix PB_1 are null and the q last rows give the identity matrix I_q of order q. Set $\hat{B} = PB_1$ and $\hat{A} = PAP^T$; then condition (E_1) is equivalent to the condition

(E_1') $y^T\hat{A}y > 0$ for all $y \in R^n$ such that $y \neq 0$ and $\hat{B}^Ty = 0$,

and condition (C_1) is equivalent to the condition

(C_1') $y^T\hat{A}y \geqslant 0$ for all $y \in R^n$ such that $\hat{B}^Ty = 0$.

Define

$$\hat{\mathscr{A}} = \begin{pmatrix} P & 0 \\ 0 & I_q \end{pmatrix} \begin{pmatrix} A & B_1 \\ B_1^T & 0 \end{pmatrix} \begin{pmatrix} P^T & 0 \\ 0 & I_q \end{pmatrix};$$

then $\mathrm{In}(\mathscr{A}_1) = \mathrm{In}(\hat{\mathscr{A}})$. The matrix $\hat{\mathscr{A}}$ can be partitioned as

$$\hat{\mathscr{A}} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} & 0 \\ \hat{A}_{12}^T & \hat{A}_{22} & I_q \\ 0 & I_q & 0 \end{bmatrix}.$$

Clearly, condition (E_1') is equivalent to the positive definiteness of the matrix $\hat{A}_{11}$, and condition (C_1') is equivalent to the positive semidefiniteness of the same matrix. The matrix

$$M=\begin{bmatrix}\hat{A}_{22} & I_q\\ I_q & 0\end{bmatrix}$$

is nonsingular, and $\mathrm{In}(M)=(q,q,0)$. See for instance Lemma 1 in Haynsworth and Ostrowski [9] or Cottle [3]. For a direct and very short proof consider the matrix

$$M_t=\begin{bmatrix}\hat{A}_{22}+tI_q & I_q\\ I_q & 0\end{bmatrix}, \qquad t\in R;$$

then M_t is nonsingular for all t. Hence $\pi(M_t)$ and $\nu(M_t)$ are not dependent on t. Letting $t\to-\infty$ and $t\to+\infty$, we obtain $\pi(M_t)\geqslant q$ and $\nu(M_t)\geqslant q$ respectively.

Now, it is easily seen that $\hat{\mathscr{A}}/M=\hat{A}_{11}$. Hence, $\mathrm{In}(\hat{\mathscr{A}})=\mathrm{In}(M)+\mathrm{In}(\hat{\mathscr{A}}/M)$. Thus we have obtained the following result.

THEOREM 1. *The bordered matrix $\mathscr{A}$ has at least q positive, q negative and $p-q$ zero eigenvalues. Furthermore, condition* (E_1) *is equivalent to the condition*

(E_2) *$\mathscr{A}$ has exactly n positive eigenvalues*

and condition (C_1) *is equivalent to the condition*

(C_2) *$\mathscr{A}$ has exactly q negative eigenvalues.*

It follows that in the particular case where $p=q$, condition (E_1) *holds if and only if $\mathscr{A}$ is not singular and condition* (C_1) *holds.*

The idea of applying the Schur complement to the study of the restriction of a quadratic form is due to Cottle [3]. The formulation of the above theorem seems to be new. (A previous version of this paper was seen by S. Schaible, and he points out a similar, but weaker result which was obtained by G. Wolkowicz [15] but not published.)

4. AUGMENTABILITY CONDITIONS

It follows from the result on the Schur complement that

$$\mathrm{In}\begin{pmatrix} A & B^T \\ B & tI_p \end{pmatrix} = \mathrm{In}(tI_p) + \mathrm{In}\left(A - \frac{1}{t}BB^T\right) \qquad \text{for all} \quad t \neq 0.$$

For the sake of simplicity we shall denote by $\mathscr{A}_t$ the first matrix in the above formula and by A_t the last one. The function $\xi(t) = \pi(\mathscr{A}_t)$ is nondecreasing and lower semicontinuous. A necessary and sufficient condition to have $\pi(\mathscr{A}) \geqslant n$ is that there exists some $\bar{t} < 0$ such that $\pi(\mathscr{A}_{\bar{t}}) \geqslant n$, or equivalently $\pi(A_{\bar{t}}) = n$, since the order of $A_{\bar{t}}$ is n. Apply Theorem 1 to obtain

THEOREM 2. *Condition* (E_1) *is equivalent to the condition*

(E_3) *there exists* $k > 0$ *such that* $A + kBB^T$ *is positive definite.*

The above result was first quoted by Finsler [6] and thereafter proved in many different ways. The proof given above seems to be one of the simpler ones.

Similarly, the function $\mu(t) = \nu(\mathscr{A}_t)$ is nonincreasing and lower semicontinuous. It follows that a necessary and sufficient condition to have $\nu(\mathscr{A}) = q$ is that there exists some $\bar{t} > 0$ such that $\nu(\mathscr{A}_t) = q$ for all $t \in (0, \bar{t}]$. Thus we have obtained an analogous theoretical result to Theorem 2, but one that is more difficult to handle.

THEOREM 3. *Condition* (C_1) *is equivalent to the condition*

(C_3) *there exists* $\bar{k} > 0$ *such that for all* $k \geqslant \bar{k}$ *the matrix* $A - kBB^T$ *has exactly* q *negative eigenvalues.*

Clearly $\nu(\mathscr{A}) \geqslant \nu(A_k) = \nu(A - kBB^T) \geqslant \nu(A)$ for all $k > 0$. Hence we deduce the following well-known result:

PROPOSITION 4. *If condition* (C_1) *holds, then* A *has at most* q *negative eigenvalues.*

5. DETERMINANTAL CONDITIONS

Equivalent conditions to conditions (C_1) and (E_1) in terms of determinants are usually obtained from Finsler's theorem. We think that a more natural way is to derive them from the conditions on the inertia of the bordered matrix. Throughout this section, we shall assume, as usually done, that $q = p$ and the $p \times p$ matrix obtained from B by keeping only the first p rows is not singular. These two conditions are not very restrictive. The first one consists of deleting the redundant equations from $B^T x = 0$, and the second amounts to performing a permutation on the coordinates of vectors $x \in R^n$.

Given $R \subset \{1,2,\dots,n\}$, we shall denote by A_R the matrix obtained from A by keeping only the rows and the columns corresponding to R; similarly, B_R is obtained from B by keeping the rows corresponding to R, and $\mathscr{A}_R$ is the matrix

$$\mathscr{A}_R = \begin{pmatrix} A_R & B_R \\ B_R^T & 0 \end{pmatrix}.$$

When $R = \{1,2,\dots,r\}$, then $A_r, B_r, \mathscr{A}_r$ stand for $A_R, B_R, \mathscr{A}_R$ respectively.

THEOREM 5 (Debreu [5], Samuelson [13]). *Assume that $p = q$ and B_p is not singular. Then condition* (E_1) *is equivalent to the condition*

(E_4) $(-1)^p \det(\mathscr{A}_r) > 0$ *for* $r = p+1,\dots,n$,

and condition (C_1) *is equivalent to the condition*

(C_4) $(-1)^p \det(\mathscr{A}_R) \geqslant 0$ *for all* $R \supset \{1,2,\dots,p\}$.

Proof. If condition (E_1) holds, then necessarily for all $R \supset \{1,2,\dots,p\}$ the condition

$$x_R^T A_R x_R > 0 \qquad \text{whenever} \quad x_R \neq 0 \text{ and } B_R^T x_R = 0$$

is satisfied. Since the rank of B_R is p, according to Theorem 1 one has $(-1)^p \det(\mathscr{A}_R) > 0$. In the same manner, if condition (C_1) is satisfied, then $(-1)^p \det(\mathscr{A}_R) \geqslant 0$.

Conversely, since B_p is not singular, then

$$\begin{pmatrix} (B_p^T)^{-1} A_p B_p & I_p \\ I_p & 0 \end{pmatrix} = \begin{pmatrix} (B_p^T)^{-1} & 0 \\ 0 & I_p \end{pmatrix} \begin{pmatrix} A_p & B_p \\ B_p & 0 \end{pmatrix} \begin{pmatrix} B_p^{-1} & 0 \\ 0 & I_p \end{pmatrix}.$$

Hence $\mathrm{In}(\mathscr{A}_p) = (p, p, 0)$ and $(-1)^p \det(\mathscr{A}_p) > 0$.

Note that there exists a finite sequence $\{R^k\}$, $k=0,\ldots,m$, such that

$$R^0 = \{1,2,\ldots,p\} \subset R^1 \subset \cdots \subset R^m \subset \{1,2,\ldots,n\},$$

$$\operatorname{card}(R^k) = p+k, \quad \det(\mathscr{A}_{R^k}) \neq 0 \qquad \text{if} \quad k=0,1,\ldots,m,$$

$$\det(\mathscr{A}_R) = 0 \qquad \text{for all} \quad R \supset R^m \text{ and } R \neq R^m.$$

Now, notice that $\mathscr{A}_{R^{s+1}}/\mathscr{A}_{R^s}$ is a 1×1 matrix and

$$\det(\mathscr{A}_{R^{s+1}}) = \det(\mathscr{A}_{R^{s+1}}/\mathscr{A}_{R^s})\det(\mathscr{A}_{R^s}),$$

$$\operatorname{In}(\mathscr{A}_{R^{s+1}}) = \operatorname{In}(\mathscr{A}_{R^{s+1}}/\mathscr{A}_{R^s}) + \operatorname{In}(\mathscr{A}_{R^s})$$

for all $s=0,1,\ldots,m$. It follows that the inertia of $\mathscr{A}$ can be easily obtained from the sequence of the signs of $\det(\mathscr{A}_{R^{s+1}})/\det(\mathscr{A}_{R^s})$. Then clearly ($E_4$) implies ($E_2$), and ($C_4$) implies ($C_2$). ∎

6. CONDITIONS IN TERMS OF ROOTS OF A POLYNOMIAL EQUATION

In this section, we also assume that $p=q$. Given $t\in R$, we define the $(n+p)\times(n+p)$ matrix

$$M_t = \begin{pmatrix} A+tI_n & B \\ B^T & 0 \end{pmatrix}.$$

Let $P(t) = \det(M_t)$. For large positive values of t the matrices $A+tI_n$ and $-A+tI_n$ are positive definite, and condition (E_1) holds when $A+tI_n$ stands for A or when $-A+tI_n$ stands for A and $-B$ for B. Hence for large positive values of t, $\operatorname{In}(M_t) = \operatorname{In}(-M_{-t}) = (n,p,0)$ [and $\operatorname{In}(M_{-t}) = (p,n,0)$]. On the other hand, if $\bar{t}$ is a root of the equation $P(t)=0$, then $\bar{t}$ is an eigenvalue of the symmetric matrix

$$\begin{pmatrix} A+2tI_n & B \\ B^t & tI_p \end{pmatrix}$$

and $\bar{t}$ is real. All the roots of the equation $P(t)=0$ are real.

The function $\pi(M_t)$ and $\nu(M_t)$ are respectively nondecreasing and nonincreasing; the jumps of these functions correspond to the roots of the equation $P(t)=0$. From Theorem 1 we obtain the following result:

THEOREM 6. *If $p=q$, condition* (E_1) *is equivalent to the condition*

(E_5) *all the roots of the equation $P(t)=0$ are negative,*

and condition (C_1) *is equivalent to the condition*

(C_5) *all the roots of the equation $P(t)=0$ are nonpositive.*

The above result was directly proved by Hancock [7], who used it to obtain some of the determinantal conditions given in Section 5.

This theorem may be used to derive another condition equivalent to condition (E_1):

THEOREM 7. *If $p=q$, condition* (E_1) *is equivalent to the condition*

(E_6) *$\mathscr{A}$ is not singular, and the $n\times n$ matrix obtained from $\mathscr{A}^{-1}$ by keeping only the first n rows and the first n columns is positive semidefinite.*

Proof. If condition (E_1) holds, then necessarily $\mathscr{A}$ is not singular. If $-k$ is a nonzero root of the equation $P(t)=0$, then there exists $(x,u)\in R^n\times R^p$, $(x,u)\neq(0,0)$, such that

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}\begin{pmatrix} x \\ u \end{pmatrix}=k\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ u \end{pmatrix}.$$

Notice that one cannot have $x=0$, because in this case $Bu=0$ is in contradiction with $u\neq 0$ and the condition on the rank of B.

Assume that $\mathscr{A}^{-1}$ is partitioned as

$$\mathscr{A}^{-1}=\begin{pmatrix} K & L \\ L^T & M \end{pmatrix}$$

where K is a $n\times n$ matrix, L is a $n\times p$ matrix, and M is a $p\times p$ matrix. Then $x=kKx$, $u=kL^Tx$. Hence the nonzero roots of the equation $P(-t)=0$ are the reciprocals of the nonzero eigenvalues of the matrix K, and the equivalence between (E_1) and (E_6) follows. ∎

7. THE PARTICULAR CASE WHERE A IS NOT SINGULAR

If A is nonsingular, then

$$\mathrm{In}(\mathscr{A}) = \mathrm{In}(A) + \mathrm{In}(\mathscr{A}/A) = \mathrm{In}(A) + \mathrm{In}(-B^T A^{-1} B).$$

The following proposition follows straightforwardly.

THEOREM 8. *If A is nonsingular, then condition* (E_1) *is equivalent to the condition*

(E_7) $\pi(A) + \nu(B^T A^{-1} B) = n$,

and condition (C_1) *is equivalent to the condition*

(C_7) $\nu(A) + \pi(B^T A^{-1} B) = q$.

When A is singular, it is possible to extend this last result by using the concept of pseudoinverse defined by Moore and Penrose. The interested reader is referred to Chabrillac [2].

REFERENCES

1 R. E. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1970.
2 Y. Chabrillac, Semi-definie positivité de formes quadratiques sur un sous espace de R^n. Fonctions monotones dans R^n, *Thèse de 3ème cycle*, Univ. de Clermont II, 1982.
3 R. W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.* 8:189–211 (1974).
4 J.-P. Crouzeix and J. A. Ferland, Criteria for quasiconvexity and pseudoconvexity, relationships and comparisons, *Math. Programming* 23:193–205 (1982).
5 G. Debreu, Definite and semidefinite quadratic forms, *Econometrica* 20:295–300 (1952).
6 P. Finsler, Über das Vorkommen definiter und semidefiniter Formen und Scharen quadratischer Formen, *Comment. Math. Helv.* 9:188–192 (1937).
7 M. Hancock, *Theory of Maxima and Minima*, Ginn, Boston, 1917; Dover, New York, 1950.
8 E. V. Haynsworth, Determination of the inertia of a partitioned hermitian matrix, *Linear Algebra Appl.* 1:73–81 (1968).
9 E. V. Haynsworth and A. M. Ostrowski, On the inertia of some classes of partitioned matrices, *Linear Algebra Appl.* 1:299–316 (1968).
10 M. R. Hestenes, Augmentability in optimization theory, *J. Optim. Theory Appl.* 32:427–440 (1980).

11 D. V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36:187–295 (1981).
12 H. B. Mann, Quadratic forms with linear constraints, *Amer. Math. Monthly* 50:430–433 (1943).
13 P. A. Samuelson, *Foundations of Economic Analysis*, Harvard U.P., Cambridge, Mass., 1947.
14 S. Schaible, Generalized convexity of quadratic functions, in *Generalized Concavity in Optimization and Economics* (S. Schaible and W. T. Ziemba, Eds., Academic, New York, 1981, pp. 183–197.
15 G. Wolkowicz, unpublished manuscript.

Received 1 November 1982; revised 26 September 1983